

# PARTICIPATION IN CONTEXT: AN EXPLORATORY STUDY OF QUERYING IN PARTICIPATORY APPLICATIONS

CAREN CROWLEY, WILFRIED DANIELS, RAFAEL BACHILLER, WOUTER  
JOOSEN AND DANIEL HUGHES

iMinds-DistriNet, KU Leuven, Celestijnenlaan 200A, Heverlee, B-3001, Belgium  
caren.crowley@cs.kuleuven.be

## ABSTRACT

Participatory applications rely upon crowd sourcing a community of users to contribute resources such as: text, audio and images. However, the success of participatory applications is dependent user responsiveness to queries. We report the results of a 30-day study examining factors affecting user responsiveness to application queries.

## INTRODUCTION

An increasing number of participatory applications are emerging which rely upon a user community to contribute content such as: text, audio, images and video (Christin, Reinhardt, Kanhere & Hollick 2011; Hoseini-Tabatabaei, Gluhak, & Tafazolli 2013). The participatory application model is *scalable*, *low cost* and enables the development of a new class of *user-centric* applications. Participatory applications can be divided into two groups, automatic and manual participatory applications. In the case of automatic participatory applications the level of human involvement required is relatively low. Data is automatically sensed by software running on the user's device and reported to the application servers examples include sensing real-time traffic conditions (Mohan, Padmanabhan, & Ramjee 2008) or human-centric activities and location (Eisenmann, Miluzzo, Peterson & Campbell 2007).

In the case of manual participatory applications, higher levels of human involvement are both enabled and required. This allows for the capture and creation of content and phenomena, which would not be possible in an automatic fashion. Examples include weather reports (Demirbas, Bayir, Akcora, Yilmaz & Ferhatosmanoglu 2010) and micro-blogs involving audio and visual content (Gaonkar, Li, Choudhury, Cox & Schmidt 2008). However, the benefits of this application model must be balanced against the inherent risks of expanded human involvement, as users are inherently unreliable. In terms of user responsiveness, the risks are two-fold: Firstly, there is a risk that query messages will be lost if the user does not notice them or chooses to ignore them. Secondly, users may take an arbitrarily long time to respond to a message, which is a problem for time-sensitive applications. This poses the question: *what can be done to increase user responsiveness to application queries?*

We conceptualise a manual participatory application as a network arrangement and investigate message responsiveness as the outcome of the interaction between attributes of *queries* and attributes of *users*. The contributions of this paper are twofold. Firstly, we develop a model that relates the context of a user to their query

handling performance. Secondly, we elaborate upon the implications of our findings for researchers and practitioners working in the area of participatory applications.

## RELATED WORK

Manual participatory applications are an emerging area and much of the prior work has focused on automatic *participatory sensing* architectures. As such research has focused primarily on building appropriate software architectures (Jia, Xing, Tso & Dai, Ruzzelli, Jurdak & O'Hare 2007; Rachurri, Mascolo, Musolesi & Rentfrow 2011) and the deployment of sensing capabilities in real world settings (Milluzzo et al. 2008; Eisenmann et al. 2007). In automatic participatory sensing applications data is collected automatically by an application running on the participants' devices and transmitted to the servers of the participatory application. In contrast, in the case of manual participatory applications queries are sent from the application to the user, requiring a response from the user. Lane, Eisenmann, Musolesi, Miluzzo & Campbell (2008:12) argue that in the case of applications where users must consciously choose to answer queries, socio-technical techniques must be developed to encourage user involvement. Such techniques are necessary in order to ensure the participation of a large community of users and to help mitigate the negative impact that arises when a participatory application interrupts normal user behaviour.

### Query Attributes

Querying is a significant component of existing manual participatory applications. For instance Gaonkar et al. (2008) found querying to be a necessary feature of their Micro-Blog application. The application allows users to create multimedia blogs. Taking advantage of the embedded capabilities of smartphones users can upload blog entries involving a mix of audio, images and text content. These micro-blogs are associated with a date, time and location and positioned on a map shown in a graphical user interface. An Internet user can then zoom into the map and browse streaming content originating from their preferred region. Interestingly, while users were free to upload content according their personal preferences, the trial found that users were much more likely to upload content in response to query requests. The success or failure of this class of applications is then dependent on users responding to queries in a timely fashion. As Queries sent from the application to the user are the key governing mechanism in the network, it stands to reason then that altering the attributes of query messages is likely to influence user responsiveness to those queries. In particular, we focus on the attributes of queries that are both remotely observable and relatively easy to manipulate or monitor such as the messaging platform that is used to deliver messages, the time at which messages are sent and the rate at which users are messaged.

In terms of the communication channel through which users are contacted, Online Social Networks (OSNs) are growing in popularity as a medium to support participatory applications. This popularity is driven by two factors. Firstly, OSNs have a massive existing user population. Facebook has over 1.2 billion monthly-active users (Facebook 2012), while Twitter has over 500 million (Twitter 2012). Secondly, these platforms provide a natural mechanism through which applications can interact with users across many different devices (from desktops to mobile phones). Demirbas et al. (2010) for instance, recruits users in this manner. However, the

feasibility of using OSNs as a generic mechanism to recruit and communicate with participants requires more detailed examination.

The timing of messages and when the user receives them is also likely to influence responsiveness. For instance, research on survey response has found that surveys are more likely to be completed if they are received on a Tuesday or Wednesday afternoon relative to other days of the week (Dillman, Singh & Christian 2000). The importance of networking has also been recognised by social applications such as GroupOn who send promotional emails to their members at 10AM as this time was found to be optimal for their target audience of stay at home parents (Park & Chung 2012). The rate of querying is also important (i.e. the number of messages that are sent to a user within a given time-period). While more frequent querying may improve the freshness or spatial resolution of application data, it is intuitive that users will only be capable of handling a limited number of query messages. Lane et al. (2008) argue that the probability of human cooperation is likely to fall as the daily barrage of queries causes fatigue and eventually annoyance.

### **User Attributes**

In this paper, we examine the influence of remotely observable social factors on participation. Specifically, we focus on the factors that directly influence the decision of a participant to join the application (recruitment) and the subsequent participation of that individual in the network (responsiveness). Granovetter (1978) developed the concept of tie strength, distinguishing between strong and weak ties. We anticipate that stronger ties between participants and their recruiters will lead to greater levels of participation. In addition to tie strength, the social distance between a participant and those managing the application may influence participation (Efstratiou et al. 2012). Such concerns could be overcome when users were able to control the type of data released.

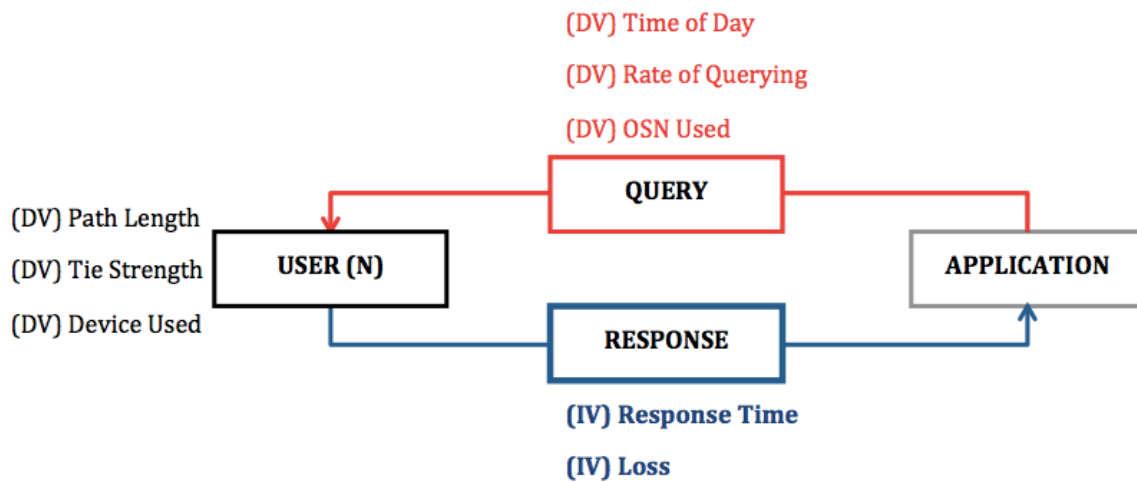
The devices through which users respond to queries may also be important. Advances in smart phone technology mean that non-professionals are likely to possess all the necessary to quickly and conveniently create, capture and report data and other content to the application (Hoseini-Tabatabaei et al., 2013). However, while smart-phones are becoming increasingly ubiquitous, the limited user interface of these devices and prohibitions against their use in the workplace may have a negative effect on the responsiveness of their users.

### **Response Attributes**

When considering user responsiveness, two attributes of response messages are critical: whether the query was lost (i.e. whether the participant fails to respond) and the latency of the response (i.e. how long the participant took to respond). The timeliness of user responses is especially critical for time sensitive applications. This can be clearly seen in the case of the *Twitter Weather Radar* application (Demirbas et al. 2010). The weather monitoring application achieved an average accuracy of 85% for current weather conditions. However, in terms of user participation, Demirbas et al. (2010) reported only a 15% response rate to queries and slow response times, with 50% of responses taking longer than 30 minutes. While the high accuracy of the application demonstrates the potential of participatory sensing applications, slow user response times and high loss rates are a clear pitfall of manual sensing approaches.

## Conceptual Framework

A *network perspective* provides an understanding of message responsiveness as the outcome of the combined influence of actor attributes and tie attributes. In this case, the actor types are the *participants* and *the application* (i.e. a 2-mode network). The ties, which link these actors, are directed. *Queries* are sent from the application to the participants while *responses* are sent from participants to the application. Figure 2 visualises this theoretical framework and illustrates how this network operates along with the relevant Dependent Variables (DV) and Independent Variables (IV). The precise definition of the variables and how they were measured is provided in the Section 3.4



**Figure 2: Theoretical Framework - A Network Perspective on User Responsiveness**

## METHODOLOGY

In order to examine how contextual factors affect message loss and user responsiveness, we designed an original experiment that ran for 30 days from December 15<sup>th</sup> 2012 to January 14<sup>th</sup> 2013. In total 70 participants were recruited for the study and a total of 3,055 automatic query messages were sent to users during the experiment.

Participants were enlisted using viral recruitment, which involved the sharing of short message on the Facebook and Twitter, online social networks. The message used to recruit participants was limited to 140 characters as this is the maximum size of a single Twitter message, including a HTTP link to the sign-up and information webpage.

Custom OSN client applications were created for managing the automatic messaging of clients. In the case of Twitter, the client sent a directed tweet (Twitter message) to participants, while in the case of Facebook; the client sent a private chat message to participants. Following self-registration, participants began to receive messages from the client application at a random time of day, but at a controlled rate of messaging.

The distribution of messages to users via the online social network clients was scheduled using standard Linux CRON jobs. During the 30-day experiment a total of 3,055 query messages were sent to the 70 participants in the fashion described above.

A QUERY has three important attributes, which can be controlled by the APPLICATION: the *time-of-day* when the user is queried, the *rate of querying* and the *OSN used* to query the user. Each USER has three important attributes: their social distance or *path length* from the application managers, the *tie strength* between the user and their recruiter and the *device used* to access the OSN. While an application cannot control the attributes of a user, it can use these attributes to inform which users are selected from the population to receive a particular query.

## RESULTS

### Message Loss

We performed logistic regression analysis to examine the effect of relevant predictor variables, time of day, online social network, rate of messaging and path length on message loss. We then performed a separate sub-group analysis of the effect of tie strength on message loss. We report both the main effects and two-way interaction effects between variables.

The main effects for both *OSN* and *Rate* have a statistically significant effect on Message Loss. We find that, on average, Facebook messages have lower rates of loss compared to Twitter. However, as there is a statistically significant effect for the interaction term ***OSN\*Rate*** we need to be careful in interpreting the main effect. We find that queries delivered via Facebook are less likely to be lost when the rate of messaging is 1 or 2 messages per day. However, this advantage is lost when the rate of messaging increase to 3 messages per day and beyond.

The effect of *Time of Day* on message loss is not significant. However there is a significant interaction between *Time of day and Rate of Messaging*. We found that Rate of Messaging has the greatest impact on message loss during the *Night* and *Morning* periods.

We performed sub-group analysis on the effect of *Tie Strength* on message loss using binary logistic regression analysis. We do not include tie strength as an independent variable in the main analysis as we only have tie strength data for 70% of the participants. We find a significant negative relationship between tie strength and message loss, implying that the stronger the relationship between a user and the person who recruited them, the less likely they are to lose messages.

### Message Response Time

In order to analyse the effect of the relevant predictor variables (Time of Day, OSN, Device, Path Length) on Message Response Time we perform a four-way ANOVA. We examined both the main effects and interaction effects. However, we only interpret interactions involving two variables.

There was a significant main effect for *Time of Day* on Response Time. However we also find significant interaction between, *Time of Day and Device*. Interestingly, while

we find that cellular devices have generally lower response times, during the afternoon period non-cellular devices have lower response times. This may imply that users are more likely to access their chosen OSN and then respond to the query on their work based non-cellular device during the afternoon.

The main effect of *OSN* was not found to be significant. However, we find a significant interaction between *OSN and Rate of Messaging*. This implies that the effect of OSN on Response Times is dependent on the rate of messaging. We find that when the rate of messaging increases beyond 3 messages per day the trend reverses with messages delivered via Twitter having a lower response time.

We find a significant main effect for *Path Length*. We find that users with higher social distance from the application managers have a higher response times. It may be the case that such users have self-selected to participate in the experiment due to their interest in the experiment rather than due to social influence.

## CONCLUSIONS AND FUTURE WORK

Improving user responsiveness to queries requires full consideration of the attributes of the network within which the application is embedded, including both query attributes and user attributes. Facebook performs better than Twitter for low-rate querying of users for non-time sensitive information, however Twitter offers superior performance when higher rates of querying are necessary or the requested information is time-sensitive. Users tend to respond quicker using cellular devices in the morning and evening, while non-cellular devices perform better in the afternoon. User response times are slow for both classes of device during the night. Message loss and User Response times are not negatively effect by tie social distance between users and their recruiter (tie strength) or users and the application managers (path length). Thus, large-scale viral recruitment for participatory applications may be feasible using online social networks, even in cases where the ties between recruiters and potential users are relatively weak.

Building on this exploratory study, we intend to undertake two complementary streams of further research. Firstly, research is required to build a better understanding of how online social networks can be used to *build communities for participatory applications*. Secondly, the distinct characteristics of user responsiveness different message types motivate further research into how the characteristics of specific *communication channels impact user responsiveness* in social applications.

## REFERENCES AVAILABLE FROM THE AUTHORS